DOCUMENT RESUME

ED 212 639 SP 019 720

AUTHOR Uhlir, Ann
TITLE Student Performance Evaluation. Physical Educators for Equity. Module 7.
INSTITUTION Eastern Kentucky Univ., Richmond. Dept. of Physical Education.; Education Development Center, Inc., Newton, Mass. Women's Educational Equity Act Dissemination Center.
SPONS AGENCY Women's Educational Equity Act Program (ED), Washington, D.C.
PUB DATE 81
NOTE 35p.; For related documents, see SP 019 713-719.
AVAILABLE FROM WEEA Publishing Center, Educational Development Center, 55 Chapel Street, Newton, MA 02160 (Set of 7 modules: $8.00).

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Ability Grouping; *Achievement Rating; Equal Education; Grading; *Holistic Evaluation; Independent Study; Individual Development; Informal Assessment; Nondiscriminatory Education; *Physical Education; *Sex Fairness; Sex Stereotypes; *Student Evaluation; Teacher Education; *Tests

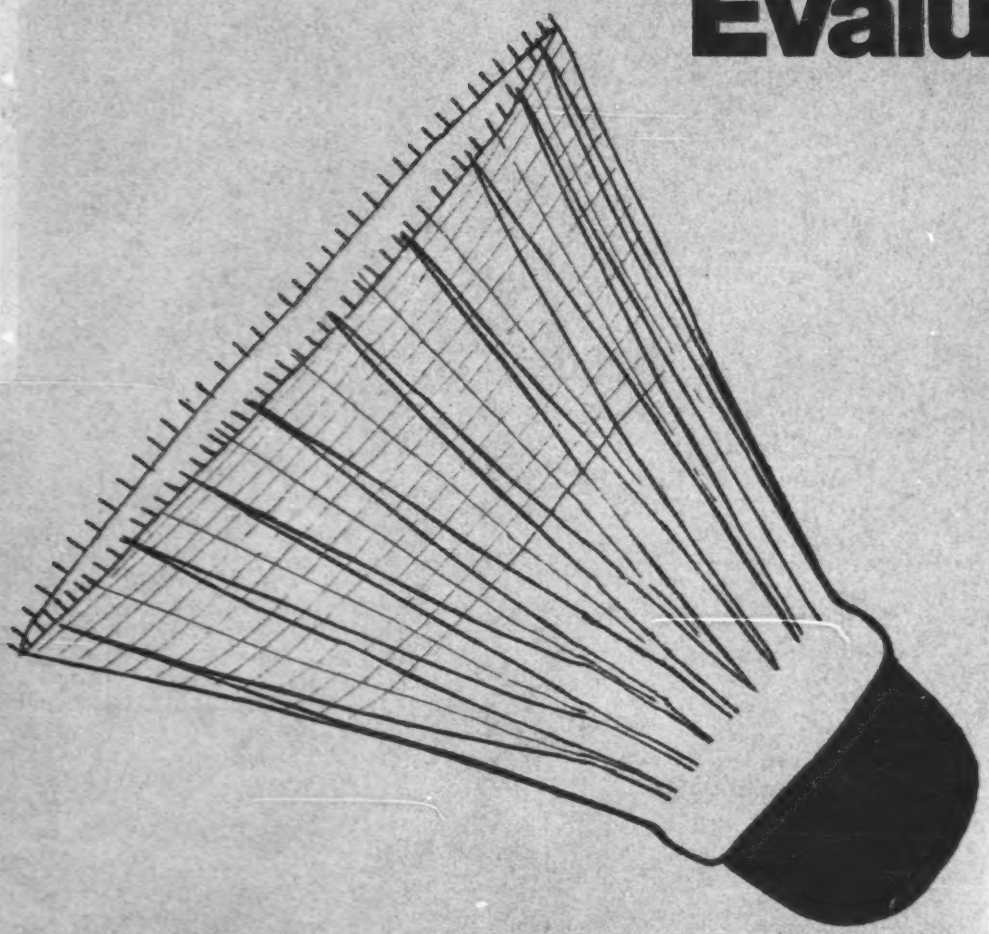
ABSTRACT

Guidelines are presented to aid secondary school physical education teachers in evaluating student performance in a way that avoids sex-role stereotyping and sex discrimination. Suggestions made for conducting testing in a bias-free setting include: (1) avoid sex-differentiated role tasks; (2) organize motor-performance testing procedures so that all test groups contain girls and boys; (3) give equal support to efforts of both sexes; (4) expect the maximum effort from all students; (5) encourage students to support the efforts of all to do as well as possible; (6) show no favoritism; (7) avoid stereotyping either sex as being superior or inferior; and (8) avoid comparing individuals of one sex with the opposite sex as a group. Guidelines are also provided to assist teachers in avoiding sex discrimination when they are grouping students by ability. A discussion is presented on the relative merits of grading by single or separate standards. Examples are given of each type of standard and the ways in which it may be used to achieve equitable results. An exploration is made of ways in which individual progress may be measured in bias-free evaluation. (JD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED212639

Student Performance Evaluation



U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION

Physical Educators for Equity

7

PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Eastern Kentucky University
Richmond, Kentucky

Women's Educational Equity Act Program
U.S. Department of Education

PHYSICAL EDUCATORS FOR EQUITY

MODULE 7

STUDENT PERFORMANCE EVALUATION

Author
Ann Uhlir

Project Director
Ann Uhlir

Department of Physical Education
Eastern Kentucky University
Richmond, Kentucky

Women's Educational Equity Act Program
U. S. DEPARTMENT OF EDUCATION
Terrel Bell, Secretary

Discrimination Prohibited: No person in the United States shall, on the grounds of race, color, or national origin, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving Federal financial assistance, or be so treated on the basis of sex under most education programs or activities receiving Federal assistance.

The activity which is the subject of this report was produced under a grant from the U.S. Department of Education, under the auspices of the Women's Educational Equity Act. Opinions expressed herein do not necessarily reflect the position or policy of the Department, and no official endorsement should be inferred.

Printed and distributed by The WEEA Publishing Center, 1981,
at Education Development Center, 55 Chapel Street,
Newton, Massachusetts 02160

MODULE 7

CONTENTS

Acknowledgments iv

Introduction v

Objectives vi

Introduction to Evaluation 1

Testing 3

Exercise 1 6

Exercise 2 6

Exercise 3 8

Grouping 9

Grouping by Ability 9

Grouping by Sex 10

Exercise 4 11

Grading 13

Single Standards 13

Separate Standards 14

Exercise 5 19

Individual Progress 23

Conclusion 23

References 25

Appendix A: Information for the Calculation of Mean and Standard Deviation Scores 27

Appendix B: The Normal Curve Method of Assigning Grades 31

Directions for module use: Read the module, following the instructions given throughout. At any time you may refer to preceding pages.

ACKNOWLEDGMENTS

Development of Modules 1-7 in their revised forms is the result of the professional contributions of many people. Appreciation is extended to all those who played a role in this unique effort to reduce sex bias in secondary physical education instruction and program operation.

The services of Agnes Chrietzberg, assessment specialist, were invaluable in all phases of the project, from conception to completion. Mary Neikirk, curriculum specialist, spearheaded the preparation of all modules and contributed enormously to the research and writing. Mary Dee Leslie, validation coordinator, assisted with many aspects of module development and validation. Gratitude is expressed to Peggy Stanaland, Harold Holmes, Geraldine Polvino, Dorothy Kirkpatrick, and Richard Lee Gentry for their assistance in module development and/or field testing. Carrie Haag was indispensable in her assistance with module development, cover designs, and manuscript preparation. Persons who served as validation subjects and those interviewed for module content selection deserve a particular word of thanks. Nancy Ferrell, illustrator; Darlene Ogden, instructional materials technologist; and Theresa Snow, graphic artist, offered specialized skills essential to the preparation of the materials. Linda Bain, Carol Stamm, and Marilyn LaPlante merit recognition for sharing their expertise as consultants in the areas of curriculum and assessment. Appreciation is extended to Carolyn Siegel, Reda Wolfinbarger, and Linda Holt for typing various drafts of the modules. For general assistance with a variety of assignments related to the project, Kristine Freck, Mary K. Osborne, and Becky Baker are acknowledged.

A special thank-you goes to colleagues around the country who assisted with the development and validation of the modules. The eight consultants who provided critical comments on the first draft of the modules were Dean Austin, Richard Jones, Doug Knox, Lois Kruger, Barb Landers, Diane O'Brien, Pauline Rossman, and Mike Swain. The seven national site coordinators who collected data for module validation were Wanda Buckley, Claire Combs, Fran Hermance, Joe Kohlmaier, Janet Koontz, Virginia Peters, and Judie Uhlir.

To all those who have supported this effort in many ways and especially the professionals who have worked toward a better understanding of the physical education needs of both women and men, special thanks and respect.

INTRODUCTION

This module is one of seven which are to be used as a self-study program. The modules are designed to promote the elimination of sex-role stereotyping and sex discrimination in secondary school physical education classes.

Each module contains written materials, illustrations, and learning exercises with directions for their use. At the end of each module, references are cited and resources for further study are provided. Completing each module will take a maximum of one hour, except for Module 1, which can be finished in less than one-half hour.

The content of the modules is as follows:

Module 1: Introduction to stereotyping and discrimination

Module 2: Sex-role stereotyping and its effects

Module 3: Biological sex differences

Module 4: Title IX

Module 5: Curriculum development

Module 6: Teacher behavior

Module 7: Student performance evaluation

In these modules, material which is quoted or drawn from a specific source is indicated by a reference in the text, such as (5) or (3, p. 113), corresponding to the numbered list of references at the end of each module.

Note: Throughout the modules, female high school students are referred to as girls and male high school students as boys. This is consistent with the designations used by professional organizations and associations which govern and promote various sports. However, it is recognized that in many parts of the country these students are referred to as women and men. Readers are encouraged to substitute the appropriate terms as necessary.

OBJECTIVES

Upon completion of this module, you will be able to:

1. Recognize that the teacher's influence during testing and measurement of physical performance is critical to the reduction of sex-role stereotyping in physical education.
2. Identify techniques that can be used to group students for testing and instruction.
3. Recognize the need to use separate standards for grading when either sex is being discriminated against by the use of a single standard.
4. Identify techniques that can be used to create separate standards for grading.

INTRODUCTION TO EVALUATION

Sex-role stereotyping and/or sex discrimination may occur when a physical education teacher evaluates[1] students' motor performance.[2] In order to reduce or eliminate sex discrimination and stereotyping, the teacher must attempt to create a bias-free evaluative process. Three assessment[3] duties commonly performed by the teacher are (a) creating a testing or measurement situation, (b) determining groups for instruction, and (c) assigning grades. On the following pages, bias-free approaches to testing, grouping, and grading are given.

---

[1]Evaluate (v.): make a judgment about data obtained by testing or measurement.

[2]Motor performance (n.): the demonstration of physical ability as measured by sports-skills, physical-fitness, or motor-ability tests; dance or gymnastic performance; the playing of a game; or the performing of specified skills.

[3]Assessment (n.): the process of measurement and evaluation.

TESTING

In administering a skills test or obtaining a measurement of motor performance, a teacher rarely discriminates between boys and girls in the sense that one sex is favored over the other. However, sex-role stereotyping may occur in a testing situation.

Below are some examples of what teachers may do that perpetuates sex-role stereotyping during testing or measuring motor performance:

1. Expect boys to perform better than girls.
2. Praise boys more than girls.
3. Encourage boys to improve, but appear not to care if girls improve.
4. Ridicule boys for not performing well, but ridicule girls for performing well, particularly if they perform better than boys.
5. Reinforce boys' performance, but reinforce girls' socially-accepted behavior.

Boys are socialized to view a test which involves physical activity as a challenge and as an opportunity to demonstrate physical prowess. This socialization process continues in physical education as boys are expected to perform well on sports-skills and physical-fitness tests. Boys are encouraged to excel and are praised and rewarded by their peers and teachers when they do so.

On the other hand, girls have had fewer opportunities than boys to be challenged or to challenge themselves physically. Few people expect girls to be highly skilled in physical activities, so girls rarely expect that of themselves. Girls are not encouraged to demonstrate competence, except in those activities which are traditionally thought of as "feminine," such as dance, gymnastics, or synchronized swimming. In particular, girls are discouraged from excelling if that means performing better than boys.

With social pressures such as those described above, in a physical education testing situation, girls may view the tests as something merely to get through, not as a challenge to show how well they can perform. Particularly in coeducational classes, a girl might be reluctant to excel if that means outperforming a boy.

Given these factors, what can be done to minimize sex-role stereotyping in the testing setting?

1. Avoid sex-differentiated role tasks. Use both boys and girls in all aspects of leadership, including:
   a. Scoring or timing
   b. Marking the court and otherwise preparing the test setting
   c. Demonstrating the test routine
   d. Officiating the test (judging whether trials are acceptable, determining loss of points on errors, etc.)

2. When motor-performance tests are identical for boys and girls, organize the testing procedures so that all test groups contain girls and boys. The mixing of boys and girls should be encouraged. Grouping which calls attention to sex reinforces and supports the differential socialization of boys and girls.

   a. Station method (when there is a battery of different tests administered to several persons concurrently):

      (1) Group students by ability.

      (2) Group students alphabetically.

      (3) Group students by height (or another physical characteristic which is a factor in performance outcome).

   b. Rotation method (when students take turns being tested at one test station):

      (1) Rotate students so that boys and girls are tested continuously (not first the girls, then the boys).

      (2) Group students for testing alphabetically.

      (3) Order the rotation according to anticipated performance, so that repeated changes of the test set-up are not required (for example, the height of the cross bar moved up and down; assistants fielding softball throws moved in and out).

   c. Partner method (when students, in pairs, test one another):

      (1) Assign partners so that boy/girl partners, boy/boy partners, and girl/girl partners measure each other whenever the height, weight, or another physical factor of one student is not essential for the maximum performance of the other student.

      (2) Pair students by height, weight, or by other variables when such variables are critical to test performance. For example, if one partner holds the feet while the other partner does bent-leg sit-ups, using partners of approximately equal weights will be necessary.

3. Equitably support the efforts of girls and boys. Give equal encouragement to excel, and give equal reinforcement of superior performance.

4. Expect the maximum effort from all students. Tell students to try to do their very best. Avoid differential expectations which presume boys should work harder.

5. Encourage students to support the efforts of all to do as well as possible. Reinforce student behavior that exhibits reassurance to other students (e.g., "Good try, Mac," or "On your next trial, Sue, try to keep closer to the wall so you'll shorten the time the ball is in the air between hits"). Correct students who make thoughtless comments to other students (e.g., "Jim, I thought you could do better than that," or "No wonder you scored so low, Fatso," or "Wow, Mary, you are good, for a girl").

6. Show no favoritism to boys as a group, to girls as a group, or to individuals.

   a. Examples of favoritism:

      (1) Girls are given one more trial than boys are.

      (2) Boys are permitted to omit certain test items in which girls generally outscore boys.

      (3) Members of athletic teams are exempt from tests.

      (4) Boys are ignored when they shoot baskets while testing is in progress; girls are asked to return to the line for their turns when they begin throwing balls to one another.

   b. Examples in which different treatment does not mean favoritism:

      (1) Handicapped persons are permitted to substitute test items when the handicapping condition is related to performance. However, a child with a clubfoot could be expected to take the standard test of arm strength.

      (2) Different standards for evaluating high-jump performance are used for boys and girls.

7. Avoid stereotyping either sex as being superior or inferior. For example:

   a. "Boys, you will probably do much better on this test than the girls."

   b. "Do the best you can, girls. If you have trouble, one of the boys will help you."

8. Avoid comparing individuals of one sex with the opposite sex as a group. For example:

   a. "Bob, you're as weak as the girls."

   b. "Sue, you run like a boy."

   c. "Don't be a sissy, Jim. Try it."

EXERCISE 1

In the following situation, determine which alternative would be the most satisfactory in achieving a desirable testing environment. Place a check-mark next to your choice.

Mr. Merry is administering the soccer wall volley to his ninth-grade students. The test consists of scoring the number of times the ball played legally hits the wall above a given line in a 60-second period. Six test stations are set up consecutively on one long wall. Students are organized so that:

_____ 1. Boys are placed at test stations 1, 3, and 5. Girls are stationed at 2, 4, and 6.

_____ 2. Students are arranged alphabetically so that those whose names begin with A-D go to station 1, E-K, station 2, etc.

_____ 3. Students are allowed to choose a test station.

ANSWERS

1. Choice 1 is incorrect. The grouping of boys together and girls together draws unnecessary attention to the sex of the students and supports sex-stereotyped performance.

2. The correct answer is 2. The alphabetical grouping permits the distribution of boys and girls in a random manner throughout all groups. There is no advantage to ability grouping for the soccer wall volley test, since each student uses equal time despite her/his ability.

3. Choice 3 is not the most satisfactory answer. Students generally choose to work with friends of the same sex. Teachers can avoid sex segregation through preliminary planning.

EXERCISE 2

A batter of tests is administered to tenth-grade students for the purpose of grouping students for instruction in assorted motor and fitness activities. Ms. Hope has planned the experience so that all students are involved in the process of administering the tests. Check any of the following plans she has which are free of sex-role stereotyping and consistent with sound assessment theory.

_____ 1. Students are grouped so that 50-yard-dash heats are based on students' best times, with the six best times in one heat, the next six times in a second heat established in a ninth-grade class, etc.

_____ 2. Boys are used for starting, and girls for timing, the 50-yard dash.

_____ 3. Boys and girls are tested on pull-ups and are evaluated by use of the AAHPER norms. The norms are available for boys only.

_____ 4. Partners for sit-ups (flexed leg) are assigned by weight.

_____ 5. Volunteers are solicited to work during a study hall to set up markings for testing the standing broad jump and the shuttle run.

_____ 6. The 12-minute run/walk is administered in two groups. Students are assigned a partner of the opposite sex. Boys run first, while girls note their distance at the end of 12 minutes. Girls run next, while their partners note the distance covered at the 12-minute whistle.

_____ 7. Girls are given three trials on the standing broad jump. In the interest of time, the number of trials for boys is reduced to two. The best trial score is recorded.

_____ 8. Boys are asked to dismantle the pull-up bars at the end of class.

ANSWERS

The correct choices are 1, 4, and 5, for the following reasons:

1. Grouping by ability for the 50-yard dash will encourage maximum performance. Although more boys may be in the faster heats and more girls in the slower heats, sex is not a factor in assignment to the heats.
4. Matching partners by weight is a good test procedure for sit-ups, since one partner must stabilize the other. Sex is not a factor in the assignment of partners.
5. Learning to set up a test may be a valuable experience for some students. Sex is not a factor in soliciting volunteers.

Choices 2, 3, 6, 7, and 8 are incorrect, because:

2. Boys and girls have sex-differentiated leadership roles.
3. Norms on the pull-up test are provided only for boys. It is not appropriate to evaluate the performance of girls on the pull-up test using those norms.
6. Girls run as a group; boys run as a group. Such grouping may not adequately challenge all students and could lead to different performance expectations by students and instructors.
7. Girls are favored by their being able to have an additional trial and therefore having one more score for consideration.
8. Dismantling equipment is assigned to boys and the interpretation is that such tasks (mechanically difficult) are more appropriate for one sex than for the other.

EXERCISE 3

Check the following statements, made by teachers during the course of testing, that reinforce sex-role stereotyping.

_____ 1. "Bill, do you know that half the girls scored more goals than you did?"

_____ 2. "The best score made on this tennis-drive test in all the tennis classes was that scored by Sally. Let's give Sally a big round of applause."

_____ 3. "Sue, I know you can do better on this test. I have seen you play and this score does not reflect your ability. What's the problem?"

_____ 4. "Girls, you really did well for a change. I'm glad you are finally trying."

_____ 5. "What's the matter, Sam? Are you afraid of sweat? The girls are working harder than you are."

_____ 6. "Tom, that was a decent gesture you made in offering to put up the net for Mary. But try to accept the fact that Mary wants to learn to do it herself."

_____ 7. "Betty, crying will not help. Why don't you want to take the test?"

_____ 8. "Boys, let's run a half-mile warm-up before we begin the test. Girls, you may run a half mile, too, if you want to, but be sure to do at least a quarter mile."

ANSWERS

The correct choices are 1, 4, 5, and 8, for the following reasons:

1. Bill is being reminded that the girls are better than he--as if that should not be the case.
4. The girls are being told that they usually do not try. The assumption is that because a person is a member of one sex group and does not try, members of the sex group as a whole do not try.
5. Sam is being told that there is something out of order when girls work harder than boys--that Sam, because he is a boy, should work harder than all girls.
8. The girls are not expected to run a half mile to warm up--just because they are girls.

Statements 2, 3, 6, and 7 do not reinforce sex-role stereotyping and therefore should not have been checked off.

GROUPING

By measuring the ability of students, teachers can gain information that is useful for the purpose of grouping for instruction. The process of ability grouping based on objective standards of individual performance is permissible under Title IX and is a sound teaching procedure.

In the past, when physical education classes were separated by sex, girls and boys were often taught different activities; for example, girls could not learn wrestling and boys could not learn field hockey. This was discriminatory because members of one sex were denied opportunities for participation in activities designated "for boys only" or "for girls only."

With the passage of Title IX and the trend toward coeducational physical education classes, students are rarely denied opportunities to learn an activity because of their sex. However, there may be some instances of discrimination which occur as a result of ability grouping for instruction. For example, one of the prerequisites for entrance into an advanced swimming class might be to swim 100 yards of the crawl in a given amount of time. Boys at advanced levels of swimming usually swim faster than girls. Setting the time for the 100-yard swim at a level at which only boys could qualify for the class would discriminate against those girls among the swimmers in the group who swim more slowly than the designated time allows. A possible solution to this problem would be to establish different qualifying times for girls and boys.

GROUPING BY ABILITY

Guidelines have been developed to assist teachers in avoiding sex discrimination when they are grouping students by ability (2, p. 12). These guidelines are:

1. Ability should be assessed for each particular unit, sport, or activity undertaken. Use of a single-ability standard for grouping in all activities could function to discriminate against students of one sex.

   Example: Students should be grouped for participation in track-and-field activities according to their abilities in this area, rather than according to their abilities as assessed in wrestling or swimming, or by some general measure.

2. Ability must be assessed by objective measures applied objectively. An objective measure is one in which two persons evaluating the same student reach the same rating. Relying on quantitative (numerical) measures is a way of increasing objectivity.

   Example: Teachers might measure the power and efficiency of swim strokes by counting the number of strokes needed to swim a certain distance, rather than by subjectively rating the strokes by observation.

3. Ability measures used in grouping should involve skills that are necessary for the activity or skills that maximize safe performance of the activity. Screening should be based on these relevant skills, rather than on sex.

   Example: The ability to perform a minimum number of pull-ups (an activity in which most females perform less well than most males) should not be a prerequisite for admission to an advanced basketball class. Arm strength has not been demonstrated to be a valid predictor of safety or performance in basketball; such a requirement would serve to reduce the number of females eligible for participation.

   Example: The ability to perform specified elementary dives could be a prerequisite for an intermediate diving class, regardless of the effect on the numbers of males and females in the class. Performance of the elementary dives is demonstrably related to student safety in intermediate diving activities.

GROUPING BY SEX

Sex-based grouping for instruction should be limited to participation in contact sports. Such grouping is not mandatory, but Title IX provides this alternative to allow for sports instruction involving physical contact. The potential for injury may be reduced by grouping. Furthermore, sports rules for some contact sports differ significantly for men and for women. Lacrosse, played by women, allows minimum contact; however, significant contact occurs when lacrosse is played by men. The educational setting can provide opportunities for such sports to be learned in accordance with anticipated future play.

Some principles to guide the teacher when he or she is grouping by sex are:

1. Students can be separated by sex only during contact-sports units.

   Example: A basketball unit is part of a curriculum including units in tennis, gymnastics, and volleyball. During the basketball unit, boys and girls may be separated with no objective test of ability having been given. During the tennis, gymnastics, and volleyball units, however, grouping would have to be based on ability.

2. Students may, in some instances, benefit from contact sports when they are participating with members of the opposite sex.

   Example: Lifesaving activities require the capacity to effectively rescue persons of diverse weights and sizes. Getting experience with swimming rescues of persons larger than oneself is critical.

3. Contact sports may be taught in a coeducational fashion, with appropriate organization precluding contact between persons of the opposite sex.

   Example: Basketball fundamentals may well be taught using drills and exercises in which contact between opposite sexes is avoided. A six-player game involving two zones (the forerunner of women's basketball) can be used. Boys can play on one end of the court, girls on the other, with changes at the quarter so that all play both defense and offense.

EXERCISE 4

Check each of the following grouping procedures that are associated with sex discrimination.

_____ 1. On the first day of class, two instructors decide to divide their group into two sections. They instruct all students born in January through June to meet with Mr. Merit and all students born in July through December to meet with Ms. Able.

_____ 2. For instructional purposes, the students are classified at the start of the semester into beginning, intermediate, and advanced groups, based on a general motor-ability test. Students remain in these groups for instruction in soccer, volleyball, badminton, and tumbling.

_____ 3. After a week of daily classes in softball skills, the instructor divides the group into four teams for a week of intra-class games. The instructor, during a planning period, places all the students on teams based on the instructor's judgment of student ability. Two teams have the better-skilled players balanced between them. The other two teams have the less-skilled players.

_____ 4. A volleyball class is to be divided so that one group contains those students having more potential and the other group contains the remaining students. The instructors use the standing vertical jump and a volleying test for classification purposes.

_____ 5. A lacrosse class is divided so that boys and girls are separated for instruction and participation without any effort having been made to classify the skills of boys or girls.

_____ 6. In order to be admitted to an elective physical education class, students are required to score above the 50th percentile of the combined norms for boys and girls on all of the AAHPER Youth Fitness Test items.

_____ 7. Students who have completed a tumbling class with a grade of B or better are permitted to participate in a special non-credit lunch-hour class in trampoline.

_____ 8. On the basis of skills tests, students are grouped for the first unit of instruction in tennis. Due to unusually frequent rain, the unit extends later than planned. The instructor, in the interest of time, opts to omit further grouping tests as the class progresses to units in golf, archery, and fencing.

ANSWERS

Items 2, 3, 6, and 8 should be checked. All other choices indicate that procedures for grouping are consistent with the guidelines designed to avoid sex discrimination. The reasons that 2, 3, 6, and 8 are counter to the guidelines are:

2. Ability must be assessed for each particular unit, sport or activity. The use of general motor-ability testing as a way of predicting a wide variety of motor abilities has not produced satisfactory validity. Specific motor-skills tests would be needed in each activity unit.

3. Ability must be assessed by objective measures. The instructor's observations of performance, depending on memory recall, is not a satisfactory procedure.

6. The AAHPER Youth Fitness Test combined norms move those girls who are at the 50th percentile among girls to lower percentiles. Some boys who are below the 50th percentile for boys move above the 50th percentile when norms are combined. The procedure used in this example limits the opportunities of girls.

8. Ability should be assessed for each activity. Despite the administrative problem involved in the loss of class time, students must have an opportunity to be grouped by ability each time a new activity is presented. An alternative would be to dissolve the original groups and continue the class without ability grouping.

GRADING

When teachers are assigning grades based on motor performance, they should take care to avoid sex discrimination. Evaluation of students' skills or progress in physical education should be based on standards which are consistent with the objectives of the course or activity. As mandated by Title IX, these standards must not have an adverse impact or a limiting effect upon students of one sex; in other words, the standards may not be discriminatory. If a limiting effect develops, one of two alternatives must be implemented: (a) separate standards must be used or (b) a standard which measures individual progress must be used.

SINGLE STANDARDS

In developing separate standards, teachers should first be sure that separate standards are actually needed. In some activities, such as folk dance or archery at short distances, girls' and boys' scores will be very similar, and neither sex would be discriminated against by having a single standard. Beginning-level activities in which strength is not a factor may not require evaluation strategies that consider boys and girls as separate groups. If students are grouped by ability for instruction, there is also the potential for use of a single grading standard. The possibility for using a single grading standard should always be explored before assuming that a need for separate standards exists.

Examples below and on the next page illustrate the use of a single standard in grading.

Example 1

At the end of a beginning archery unit taught to seventh-grade students, scores on five ends of arrows (six arrows per end) shot at a distance of 15 yards were used to measure achievement. The mean score[1] for boys was 152; the standard deviation[2] was 12. The mean score for girls was 150; the standard deviation, 11. Because the mean scores and the standard deviations were very close to the same for boys and girls, grades were assigned using a single standard.

---

[1]Mean score ($\bar{X}$) refers to the average score. Scores of all students in a given group are added together and the sum is divided by the number of scores. For information on computation of the mean, see Appendix A.

[2]Standard deviation (SD) refers to the average spread (or deviation) of a set of scores about its mean. The larger the standard deviation, the more heterogeneous the group; the smaller the standard deviation, the more homogeneous the group. For information on computation of the standard deviation, see Appendix A.

In some cases, teachers may be uncertain as to whether or not differences between mean scores for boys and girls are great enough to require using separate standards. The teacher may look for additional evidence by comparing the proportion of girls above the mean and the proportion of boys above the mean. This will help the teacher to judge whether or not using a single standard would be to the disadvantage of one sex more than to the other.

Example 2

A pretest for enrollment in swimming classes was administered to eighth-grade physical education students. The students were grouped in classes of beginning, advanced-beginning, intermediate, and advanced swimming. Minimum competencies for each letter grade were established for each class. Achievement was related to distance, form, stroke knowledge, and efficiency. Boys and girls were both graded according to these established minimums, due to the homogeneous grouping and the absence of any discernible performance differences.

SEPARATE STANDARDS

In activities in which strength, speed, power, and endurance are important factors in performance, scores on motor-performance tests are likely to be noticeably different for girls and boys. Even in activities in which physical characteristics are not influential in determining success of performance, boys often have an advantage over girls due to boys' previous experience, and thus boys often have better scores. As girls and women have more opportunities to gain physical skills and are encouraged to excel in physical activities, the need for having separate standards for males and females will probably diminish.

The following examples show how and why teachers develop separate standards for grading. The following examples illustrate the rank-order, normal curve, and T-scale methods of developing separate standards for grading.

Example 1

Rank order is a norm-referenced method of grading in which one's position within the group indicates one's grade. Students are compared with one another, rather than with a standard. If there are fifty students in the class--twenty girls and thirty boys--and the instructor has observed that the test indicates that the performances of girls surpass the performances of boys, the instructor may proceed according to the following steps:

1. Determine the percentage of each sex to receive each letter grade.
2. Rank order the grades by sex.
3. Assign grades.

The instructor should decide on the grade distribution as the following example illustrates:

A = 20% B = 30% C = 40% D = 10%

Grade	Percent Receiving Grade	Number of Girls	Number of Girls Receiving Grade	Number of Boys	Number of Boys Receiving Grade
A	20%	20	4	30	6
B	30%	20	6	30	9
C	40%	20	8	30	12
D	10%	20	2	30	3

The following tables represent actual test scores on a rhythmic rope-jumping test. Boys and girls were graded based on separate standards.

Boys

Score	Frequency	Grade
70	1	
68	1	
64	2	A
63	1	
62	1	
60	2	
59	4	B
56	3	
53	4	
50	3	
45	2	C
40	2	
36	1	
30	1	
29	1	D
25	1	
N = 30		

Girls

Score	Frequency	Grade
78	1	
76	1	A
72	1	
70	1	
65	2	
64	1	B
63	2	
62	1	
60	2	
59	3	
56	1	C
53	1	
50	1	
45	1	D
40	1	
N = 20		

Example 2

The normal curve method of assigning grades assumes that the scores are normally distributed, with most scores falling around and near the mean and with few extremely high or extremely low scores. Baumgartner and Jackson list the steps required to use this method (1, pp. 326-328). Appendix B gives an explanation of how the normal curve method can be applied to assigning grades. The following example shows the grade ranges that result from the use of the normal curve method when separate standards are indicated.

At the end of a bowling course, the teacher used students' final three game averages to establish norms as shown in the table below.

Norms	Boys	Girls
Superior	142 or up	122 or up
Good	121 - 141	100 - 121
Average	100 - 120	78 - 99
Poor	79 - 99	56 - 77
Inferior	78 or below	55 or below
Means	$\bar{X}$ = 110	$\bar{X}$ = 89

The boys' average or mean ($\bar{X}$) score is higher than the girls' mean score. Had the teacher not created separate standards for boys and girls but instead used norms based on the boys' scores, the highest-scoring girls would have been designated as average. The categories of norms could be designated as grades, so that superior = A, good = B, average = C, poor = D, and inferior = F. Then, a girl with a score of 125 would receive an A, and a boy with the same score would receive a B.

Example 3

When boys and girls must be graded by use of separate standards, the T-scale method is appropriate. A T-scale is a scale that allows one set of raw scores[1] to be converted for comparison with other sets of raw scores. A T-score of 50 would reflect that the student is at the group mean. Therefore, a T-score of 50 for Bill would mean that Bill had an average boys' score; a T-score of 50 for Bonnie would mean that Bonnie had an average score among girls. The conversion of raw scores to T-scores is particularly helpful when a battery of tests is used. T-scores for

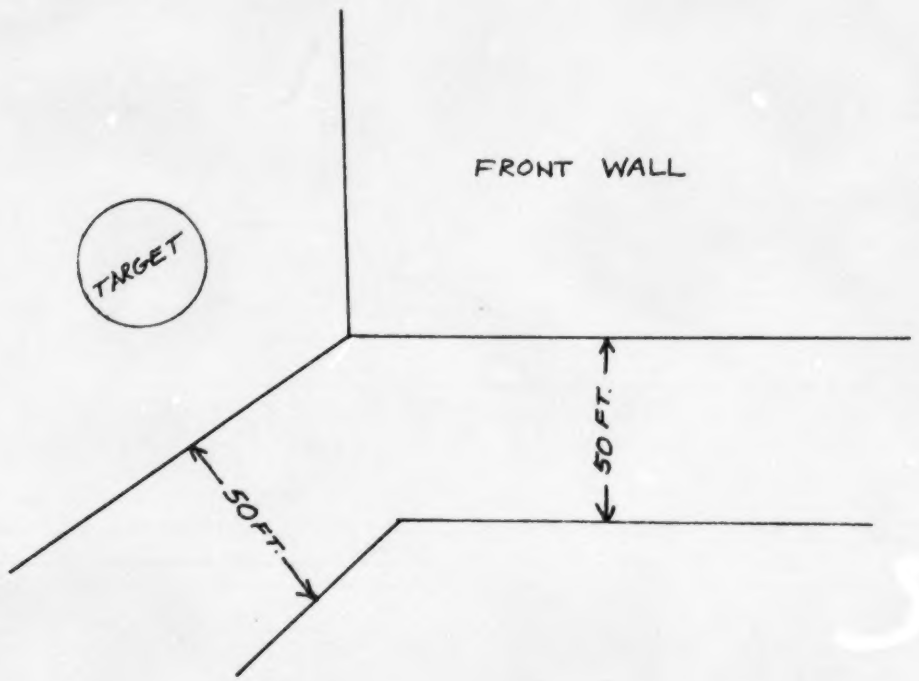
---

[1] Raw score (n.): The actual score the student receives as a result of a measurement. For example, raw scores can appear as distances, times, weights, or frequency counts.

several tests can be averaged for each sex in order to provide a composite score to serve as the basis for assigning a grade. The following example illustrates the use of the T-scale method to grade boys and girls using separate standards.

In a given teacher-made, motor-performance test for football, the student threw a ball to a front wall from a distance of 60 feet. The student then fielded the ball as it rebounded and threw it toward a target on a right-angled wall from behind a 60-foot line. The student's score was in tenths of seconds from the time of the initial throw to the time the ball hit the target. Five trials were given to each student. The sum of the two best trials was recorded. If a throw missed the target, 2 seconds were added to the time for that attempt.

A diagram of the test area appears below.



The instructor noted that times for boys and girls were different. The mean throwing time for boys' best throws was considerably higher than that for girls. The instructor therefore calculated the T-scores for boys' throws and the T-scores for girls' throws. This required first calculating the mean and the standard deviation for boys and the mean and the standard deviation for girls.

$$\text{T-score} = 50 + \frac{10\ (\text{Raw Score} - \text{Mean})}{\text{Standard Deviation}}$$

Below is a chart which shows the T-scores which were calculated for boys and for girls.

T-Scale for Softball-Fielding Test

T-Score	Time in Seconds	
	Boys	Girls
80	3.2	6.0
75	4.3	7.0
70	5.4	8.0
65	6.3	9.0
60	7.2	10.0
55	8.1	11.0
50	9.0	12.0
45	9.9	13.0
40	10.8	14.0
30	11.7	15.0
25	12.6	16.0
20	14.4	18.0
	$\bar{X}$ = 9.0 SD = 1.8	$\bar{X}$ = 12.0 SD = 2.0

Note: Data are exemplary only and do not represent actual scores.

The instructor assigned grades to students using the following grade distribution. Because T-scores were calculated for boys and girls separately, a proportionately equal number of boys and girls received the same grade.

Boys		Girls	
T-Score	Grade	T-Score	Grade
above 59	A	40 - 47	C
58 - 59	A-	38 - 39	C-
56 - 57	B+	36 - 37	D+
50 - 55	B	30 - 35	D
48 - 49	B-	28 - 29	D-
		below 28	F

For additional information on the use of the T-scale method, see Baumgartner and Jackson (1, pp. 43-47).

EXERCISE 5

For each of the following situations, place a checkmark in the space provided if that statement indicates a practice that has a strong potential for representing sex discrimination in grading. Assume that all groups are coeducational.

_____ 1. One standard is used for grading swimming form for the crawl, breaststroke, and sidestroke.

_____ 2. Students are given the option of being evaluated on four gymnastics events of their choosing. The following represent the seven selection alternatives: parallel bars, uneven parallel bars, vaulting, balance beam, still rings, free exercise, and high bar. One compulsory routine keeping sex-differentiated advantages minimal is developed in each unit. One standard is used for grading students on the compulsory routine in each of the seven events.

_____ 3. Students are given a 30-second shooting test (the number of goals scored in 30 seconds) of basketball ability. One standard is used for grading all students.

_____ 4. The instructor determines in advance that performance scores on a putting test in golf will not be subject to separate grading standards.

_____ 5. Prior to a bowling class, students are tested to determine their initial level of ability. At the end of the course, a final three-game average is used to determine achievement level. The gross difference in progress from the initial performance to the terminal performance is the basis for grading all students.

_____ 6. The instructor has designed a racquetball test and developed a T-scale for boys and a T-scale for girls. The instructor has determined that all students having a T-score below 30 will receive a failing score.

_____ 7. The instructor has constructed an intra-class tennis ladder. The students who hold the top five positions on the ladder will receive an A. The class is a beginning tennis class.

_____ 8. In a test of badminton short-serve placement, the instructor noted that the mean and standard deviation scores for boys and girls were practically the same. The instructor, therefore, decided to award grades by ranking all students together and using the following standard: A = 15%, B = 30%, C= 40%, D = 10%, F = 5%.

_____ 9. In an elective ballet class, only one boy enrolled. The instructor used one standard for grading all students. The boy received a C.

_____ 10. Weight training was one option offered in a required physical education program. Two girls enrolled. One weighed 95 pounds, and the other, 160 pounds. The instructor evaluated all students based on the percentage of body weight that could be lifted in a given number of repetitions. The curl, dead lift, press, and abdominal curl were used.

ANSWERS

Situations 3, 5, 7, and 10 should have been checked because:

3. The basketball-shooting test dependent upon vertical jump, quickness, and leg and arm strength would produce a significantly higher mean score for boys than for girls. Separate standards are necessary since the single standard has an adverse effect on girls.

5. The measurement of gross progress from initial to terminal performance not only is an unsound procedure, but also tends to favor those who had the lowest initial scores. In this case, girls would have an advantage, and the adverse impact would be on boys.

7. Even among beginning tennis players, the strength differences between boys and girls would soon be a significant variable in competition. The expectancy is that girls would be much less likely to be among the top five in the ladder tournament and would in any case represent a disproportionate percentage of the top group.

10. Weight training is highly dependent upon strength. The differences in capacities of boys and girls to develop strength is a real sex difference. Girls having weights identical to those of boys could not be expected to develop the strength required to lift (especially with their arms) a proportionately equal amount of weight. Separate standards would be required.

Situations 1, 2, 4, 6, 8, and 9 should not be checked, because:

1. Swimming form is not expected to be influenced by the sex of the student. However, there is the possibility that one sex may have had more opportunity or experience than the other to learn correct swimming form. One standard may be used if the teacher determines that there are no sex differences.

2. Allowing students the choice of any four out of seven compulsory gymnastic routines provides girls and boys with the option of selecting those events which have traditionally been taught only to one sex or only to the other sex, as well as the opportunity to be evaluated on any event the student believes to be her/his best event.

4. Putting can be expected to require a performance skill not associated with an inherent advantage of being male or female. Presuming minimal prior experience with golf, boys and girls could be expected to perform similarly on a putting test.

6. T-scores are standard scores and allow comparison.

8. When mean and standard deviation scores of two groups are practically the same, the distribution of scores within groups is very much alike. Using a single grading standard would be appropriate.

9. A teacher would not be expected to be biased against a male (or a female) student because of his or her sex. However, care must be taken to evaluate the student fairly, particularly if the boy has had less experience and opportunity than the girls enrolled in the class have had.

INDIVIDUAL PROGRESS

The progress of an individual during a period of instruction is frequently a consideration teachers wish to use as a basis for grading. The concept of basing a grade on such progress has some limitations (1, p. 320):

1. Difference scores (measuring change from an initial performance score to a final performance score) are not reliable.
2. Not all students have an equal opportunity for improvement.
3. Students are not motivated to perform well on the initial test.

Sometimes students perform poorly on a pretest or at the beginning of a unit or course so that their improvement seems greater. Students who are more skilled have less room for improvement that those who are less skilled. Thus, the raw gain scores of the former will not be as great. This condition may be found more frequently among boys with more prior skill development. Furthermore, difference scores have been notably lacking in statistical reliability.

The use of progress as a grading criterion must be pursued with caution. If, however, progress is fundamental to a teacher's basic philosophy of grading, the teacher must operate in such a way that the procedures used are as defensible as possible.

Mastery learning focuses on individual progress. Regular evaluation is used to provide feedback to students and teachers on student progress. Such evaluation of student progress is more individualized. Performance is judged against a criterion, rather than against a norm-referenced group. Evaluation of individual progress might well be used in conjunction with mastery learning to promote bias-free evaluation.

CONCLUSION

The material presented in this module should help the physical education teacher be aware that activities related to testing, grouping, and grading students in physical education should be conducted without regard to sex, except when such practices cause adverse effects for members of one sex. When real sex differences (see Module 3) exist, the teacher is obligated to follow the procedures set forth in this module or to use comparable ones which do not give an advantage to one sex at the expense of the other sex.

REFERENCES

1. Baumgartner, Ted A., and Andrew S. Jackson. Measurement for Evaluation in Physical Education. Boston: Houghton Mifflin Company, 1975.

2. Resource Center on Sex Roles in Education. Title IX and Physical Education: A Compliance Overview. Washington, D.C.: U.S. Office of Education, 1976.

APPENDIX A

INFORMATION FOR THE CALCULATION OF MEAN AND STANDARD DEVIATION SCORES

CALCULATION OF MEAN AND STANDARD DEVIATION

Formulas and simple examples for the calculation of mean and standard deviation scores are given below. Baumgartner and Jackson provide more detailed information (1, pp. 30-37).

Formula for Mean

$$\bar{X} = \frac{\Sigma X}{N}$$

Where

$\bar{X}$ = mean
Σ = sum of
X = a raw score (the actual score of a student)
N = number of scores

Example

$$\begin{array}{r} 4 \\ 8 \\ 3 \\ 4 \\ 1 \\ \hline \Sigma X = 20 \end{array}$$

$$\bar{X} = \frac{20}{5}$$

$$\bar{X} = 4$$

Formula for Standard Deviation

$$SD = \sqrt{\frac{\Sigma(X - \bar{X})^2}{N}}$$

Where

SD = standard deviation
Σ = sum of
$\bar{X}$ = mean
X = a raw score
N = number of scores

Example

X	$\bar{X}$	$(X - \bar{X})$	$(X - \bar{X})^2$
4	4	0	0
8	4	4	16
3	4	-1	1
4	4	0	0
1	4	-3	9
		$\Sigma(X-\bar{X})^2$ =	26

$$SD = \sqrt{\frac{26}{5}}$$

$$SD = \sqrt{5.2}$$

$$SD = 2.281$$

APPENDIX B

THE NORMAL CURVE METHOD OF ASSIGNING GRADES

GRADING ON THE NORMAL CURVE

SAMPLE SITUATION

In a bowling class, boys tended to score higher than girls. The final three-game average for boys was 110 and for girls, 89. Had grades been assigned using a single standard, girls would have been at a disadvantage. Separate standards for boys and girls were established. To grade on the normal curve, the following steps were followed by the instructor:

STEPS

1. Grades for boys and girls were calculated separately. Mean and standard deviation scores for each group were calculated as follows:

Boys	Girls
$\bar{X} = 110$	$\bar{X} = 89$
SD = 21	SD = 22

2. The percentage of students to receive each letter grade needed to be decided. The grade percentages shown below are frequently used to grade on the normal curve.

3. The Z-scores that correspond with the percentages are found by consulting a table showing the percentage parts of the total area under a normal curve (1, p. 51). Z-scores represent units of deviation on the normal curve.

Grade	Percentage of Students to Receive Grade	Z-Score
A	7%	+1.5
B	24%	+ .5
C	38%	- .5
D	24%	-1.5
F	7%	---

4. The cut-off points within the test scores are found by multiplying each of the normal curve values (Z-scores) by the standard deviation score of the test. Each of these values is then added to the mean score.

   Cut-off point = $\bar{X}$ + (Z-score) (standard deviation)

Boys	Girls
A = 110 + 1.5 (21) = 141.5 = 142	A = 89 + 1.5 (22) = 122
B = 110 + .5 (21) = 120.5 = 121	B = 89 + .5 (22) = 100
C = 110 + -.5 (21) = 99.5 = 100	C = 89 + -.5 (22) = 78
D = 110 + -1.5 (21) = 78.5 = 79	D = 89 + -1.5 (22) = 56

5. The cut-off values are then rounded to the nearest unit, if necessary.

6. Finally, grades are assigned, as follows:

Boys	Girls
A = 142 or up	A = 122 or up
B = 121 - 141	B = 100 - 121
C = 100 - 120	C = 78 - 99
D = 79 - 99	D = 56 - 77
F = 78 or below	F = 55 or below

☆ U.S. GOVERNMENT PRINTING OFFICE: 1981 500-226/221